Forum Statistika dan Komputasi, April 2010 p: 8-16 Vol 15 No.1
ISSN : 0853-8115

PEMODELAN RESIKO PENYAKIT KAKI GAJAH (FILARIASIS) DI PROVINSI PAPUA DENGAN REGRESI *ZERO-INFLATED* POISSON

(*The Risk of Filiarisis Disease in Papua District Modeling by Zero-Inflated Poisson*)

Sri Pingit Wulandari [1)], Brodjol Sutijo Suprih Ulama, Ika Rahmawati
Jurusan Statistika FMIPA ITS Surabaya
E-mail : [1)]sri_pingit@statistika.its.ac.id

Abstract

The goverment has established elimination of filariasis tropical disease as one of the priority programs. One of the districts that has become a target is Papua. The total amount of filariasis victim on every regency/city in Papua district can be assumed to follow a Poisson distribution. So Poisson regression method is a suitable method to know the influence factor of filariasis disease. Poisson regression model assumes equidispersion, that is equality of mean and variance of the response variable. Overdispersion test shows that the variance of the response variable exceeds its mean value. So the model is modified into zeroinflated Poisson (ZIP) regression model (logit and log). ZIP logit regression model shows that the quantity of filariasis victim in every regency/city in Papua district with zero count is influenced by the percentage of household members who sleep inside mosquito net, the percentage of household members who sleep inside insecticide musquito net, and the percentage of house-holds who keep pet (dog/cat/rabbit). While ZIP regression on log model shows that the increasing number of percentage household who keeps their pet will enhance the quantity of filariasis victim in Papua district as many as two people. Regencies/cities which need to get special attention through an elimination program of filariasis are Asmat, Tolikara, Supiori, Yapen Waropen, and Jayapura city.

Keywords : filariasis, Poisson regression, overdispersion, zero-inflated Poisson regression

PENDAHULUAN

Indonesia merupakan salah satu negara tropis yang menjadi kawasan endemik penyakit tropis, antara lain malaria, kusta, demam berdarah dengue, dan filariasis. Menurut Ambarita & Sitorus (2004), penyakit kaki gajah (filariasis) merupakan penyakit menular yang disebabkan oleh infeksi cacing filaria yang ditularkan oleh gigitan nyamuk. Penyakit ini dapat mengakibatkan menurunnya produktivitas kerja, kecacatan, stigma sosial, dan lain-lain. Pada tahun 2007 sasaran pengobatan filariasis mencapai 30 juta jiwa yang dilakukan di 72 kabupaten di Indonesia, khususnya bagian timur. Data Riset Kesehatan Dasar 2007 menunjukkan bahwa persentase penderita filariasis di Provinsi Papua berada dalam urutan tiga besar dari seluruh provinsi di Indonesia. Sehingga provinsi Papua merupakan salah satu daerah endemis yang menjadi sasaran pengobatan penyakit filariasis pada tahun 2007. Program eliminasi penyakit kaki gajah dapat dilakukan lebih efisien jika faktor-faktor yang mempengaruhinya sudah diketahui. Keterkaitan faktor-faktor tersebut dengan banyaknya penderita filaria-sis dapat didekati oleh analisis statistik mengenai hubungan variabel prediktor dengan variabel respon yaitu metode regresi. Hubungan antara penderita filariasis di Provinsi Papua dengan faktor yang mempengaruhinya dapat diketahui menggunakan metode regresi Poisson karena jum-lah penderita filariasis di Provinsi Papua sebagai variabel respon dapat diasumsikan mengikuti distribusi Poisson karena kejadian filariasis merupakan peristiwa yang relatif jarang terjadi. Menurut Khoshgoftaar, Gao, dan Szabo (2004), metode regresi mewajibkan equidispersi, yaitu nilai mean dan varians dari variabel respon harus memiliki nilai yang sama. Adakalanya terjadi fenomena overdispersi, yaitu varians dari variabel respon lebih besar dari nilai mean. Jika terjadi overdispersi, maka yang digunakan adalah metode regresi *zero-inflated* Poisson (selanjutnya disebut ZIP). Penelitian ini bertujuan untuk menganalisis karakteristik penduduk Provinsi Papua dan mengetahui faktor-faktor yang mempengaruhi kejadian penyakit tropis kaki gajah (filariasis) di Provinsi Papua.

TINJAUAN PUSTAKA

Regresi Poisson

Model regresi Poisson merupakan model standar untuk data diskrit dan termasuk dalam model regresi nonlinear (Cameron & Trivedi, 1998). Baharuddin (2005) mengatakan bahwa metode regresi Poisson biasanya diterapkan pada penelitian kesehatan masyarakat, biologi, dan teknik dimana variabel responnya (y) berupa cacahan objek yang merupakan fungsi dari sejumlah karakteristik tertentu (x). Model regresi Poisson ditulis sebagai berikut (Myers, 1990) :

$$y_i = \mu_i + \varepsilon_i = t_i \exp(x_i^T \boldsymbol{\beta}) + \varepsilon_i \quad , (i = 1, 2, \ldots, n) \quad (1)$$

dimana μ_i adalah rata-rata jumlah kejadian dalam periode t_i . Persamaan distribusi Poisson dinyatakan dengan persamaan sebagai berikut :

$$p(y_i;\boldsymbol{\beta}) = \frac{e^{-t_i[\mu(\mathbf{x}_i;\boldsymbol{\beta})]}[t_i\mu(\mathbf{x}_i;\boldsymbol{\beta})]^{y_i}}{y_i!} \quad (2)$$

dengan $\mu(x_i;\boldsymbol{\beta})$ adalah rata-rata Poisson dan $\boldsymbol{\beta}$ menunjukkan vektor parameter yang ditaksir. Selanjutnya model regresi Poisson pada Persamaan (1) dapat ditulis sebagai berikut (Myers, 1990) :

$$y_i = t_i \exp(x_i^T \boldsymbol{\beta}) + \varepsilon_i$$

Berdasarkan persamaan distribusi Poisson yang ditunjukkan pada Persamaan (2), maka fungsi kemungkinannya adalah sebagai berikut (Myers, 1990)

$$L(\mathbf{y},\boldsymbol{\beta}) = \prod_{i=1}^{n} p(y_i;\hat{\boldsymbol{\beta}}) \frac{\left\{\prod_{i=1}^{n}\left[t_i\mu(x_i;\boldsymbol{\beta})\right]^{y_i}\right\} e^{-\sum_{i=1}^{n} t_i\mu(x_i;\boldsymbol{\beta})}}{\prod_{i=1}^{n} y_i!} \quad (3)$$

Persamaan (3) dimaksimalkan dengan menggunakan teknik iteratif yang menghasilkan penaksir kemungkinan maksimum untuk koefisien regresi dalam β . Prosedur yang disarankan oleh Myers (1990) untuk menemukan penaksir kemungkinan maksimum adalah pendekatan kuadrat terkecil terboboti iteratif (*Iteratively Reweighted Least Squares*, selanjutnya disebut *IRWLS*). Menurut Cameron dan Trivedi (1998), *IRWLS* menggunakan metode Newton-Raphson. Metode ini digunakan untuk menyelesaikan persamaan berikut :

$$\frac{\partial \ln L(y;\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = 0$$

Pengujian kesesuaian model dengan devians (Kleinbaum, Kup-per, dan Muller, 1988). Berikut ini adalah hipotesis pengujian kesesuaian model regresi Poisson.

$$\mathrm{H}_0 : \mu_i = t_i\, \mu(x_i;\boldsymbol{\beta}),\ (i = 1, 2, \ldots, n)$$
$$\mathrm{H}_1 : \mu_i \neq t_i\, \mu(x_i;\boldsymbol{\beta})$$

Statistik uji yang digunakan adalah sebagai berikut

$$G = -2\ln\left[\frac{L(\mathbf{y};\hat{\boldsymbol{\beta}})}{L(\mathbf{y};\hat{\boldsymbol{\mu}})}\right] = 2\sum_{i=1}^{n}\left[y_i \ln\left(\frac{y_i}{\hat{y}_i}\right) - \left(y_i - \hat{y}_i\right)\right]$$

Daerah penolakan untuk pengujian ini adalah H_0 ditolak pada taraf signifikansi α jika $G > \chi^2_{(n-k-1,\alpha)}$, dengan n adalah jumlah pengamatan dan $k+1$ adalah jumlah parameter. Taksiran diharapkan mendekati pengamatan atau tingkat kesalahan diharapkan kecil sehingga nilai devians yang diharapkan adalah nilai devians yang kecil. Hipotesis yang digunakan untuk pengujian parameter secara individu adalah sebagai berikut :

$$\begin{array}{l}\mathrm{H}_0 : \beta_r = 0 \\ \mathrm{H}_\mathrm{i} : \beta_r \neq 0\end{array} \quad , 0 < r < k$$

dimana $k+1$ adalah jumlah parameter. Menurut Kleinbaum dkk. (1988), statistik uji yang digunakan adalah sebagai berikut :

$$G = -2\ln\left[\frac{L(y;\hat{\boldsymbol{\beta}}_r)}{L(y;\hat{\boldsymbol{\mu}})}\right] + 2\ln\left[\frac{L(y;\hat{\boldsymbol{\beta}})}{L(y;\hat{\boldsymbol{\mu}})}\right] = -2\ln\left[\frac{L(y;\hat{\boldsymbol{\beta}}_r)}{L(y;\hat{\boldsymbol{\beta}})}\right]$$

Daerah penolakan untuk pengujian ini adalah H_0 ditolak pada taraf signifikansi α jika $G > \chi^2_{(k-r,\alpha)}$ dengan $0 < r < k$.

Overdispersi

Khoshgoftaar dkk. (2004) mengatakan bahwa metode regresi Poisson mewajibkan equidispersi, yaitu kondisi dimana nilai mean dan varians dari variabel respon bernilai sama. Namun, adakalanya terjadi fenomena overdispersi dalam data yang dimodelkan dengan distribusi Poisson. Overdispersi berarti varians lebih besar daripada mean. Taksiran dispersi diukur dengan devians atau Pearson's Chi-Square yang dibagi derajat bebas. Data overdispersi jika taksiran dispersi lebih besar dari 1 dan underdispersi jika taksiran dispersi kurang dari 1.

Model Regresi *Zero-Inflated* Poisson

Jansakul dan Hinde (2001) mengatakan bahwa salah satu penyebab terjadinya over-dispersi adalah lebih banyak observasi bernilai nol daripada yang ditaksir untuk model Regresi Poisson. Salah satu metode analisis yang diusulkan untuk lebih banyak observasi bernilai nol daripada yang ditaksir adalah model regresi ZIP. Distribusi ZIP memiliki fungsi peluang sebagai berikut (Jansakul &Hinde, 2004) :

$$\Pr\left(Y_i = y_i\right) = \begin{cases} \omega_i + (1-\omega_i)e^{-\mu_i}, & y_i = 0 \\ (1-\omega_i)\dfrac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}, & y_i = 1, 2, ..., 0 \le \omega_i \le 1 \end{cases}$$

Lambert dalam Jansakul & Hinde (2001) menunjukkan model gabungan untuk $\boldsymbol{\mu}$ dan $\boldsymbol{\omega}$ sebagai berikut

$$\log(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta} \text{ dan } \operatorname{logit}(\boldsymbol{\omega}) = \log\left(\frac{\boldsymbol{\omega}}{1-\boldsymbol{\omega}}\right) = \mathbf{X}\boldsymbol{\gamma}$$

dengan $\mathbf{X}$ adalah matriks variabel prediktor, $\boldsymbol{\beta}$ dan $\boldsymbol{\gamma}$ adalah vektor parameter yang akan ditaksir, dan $\boldsymbol{\omega}$ adalah probabilitas observasi bernilai nol.

Menurut Khoshgoftaar dkk. (2004), estimasi parameter regresi ZIP dengan menggunakan metode kemungkinan maksimum. Fungsi Log-kemungkinan gabungan untuk model regresi ZIP diberikan oleh

$$\ln L(\boldsymbol{\beta}, \boldsymbol{\gamma} | y_i, x_i) = \sum_{\substack{i=1 \\ y_i=0}}^{n} \ln(\exp(x_i^{'}\boldsymbol{\gamma}) + \exp(-\exp(x_i^{'}\boldsymbol{\beta}))) - \sum_{i=1}^{n} \ln(1+\exp(x_i^{'}\boldsymbol{\gamma})) + \sum_{\substack{i=1 \\ y_i>0}}^{n} ((y_i x_i^T \boldsymbol{\beta}) - \exp(x_i^T\boldsymbol{\beta})) - \sum_{\substack{i=1 \\ y_i>0}}^{n} \ln(y_i)!$$

Estimasi kemungkinan maksimum untuk $\boldsymbol{\beta}$ dan $\boldsymbol{\gamma}$ dapat diperoleh dengan menggunakan pendekatan standard untuk model campuran, yaitu Algoritma EM. Algoritma EM memberikan prosedur sederhana yang dapat diimplementasi dalam software standar, atau metode estimasi langsung seperti metode Newton-Raphson.

Pengujian kesesuaian model dan pengujian parameter Regresi ZIP adalah dengan menggunakan uji rasio kemungkinan. Tabel hipotesis dan statistik uji untuk pengujian parameter Regresi ZIP (Lestari, 2008) tertera pada Tabel 1. Daerah penolakan untuk ketiga pengujian adalah H_0 ditolak pada taraf signifikansi α jika $G_{hitung} > \chi^2_{(v,\alpha)}$.

Pemilihan model terbaik untuk regresi ZIP, salah satunya adalah dengan metode AIC (*Akaike's Information Criterion*). Nilai AIC adalah :

$$G + (k+1)$$

dengan G adalah statistik uji kesesuaian model dan $k+1$ adalah jumlah parameter (Dalrymple *et al.* 2001). Model terbaik regresi ZIP adalah model dengan nilai AIC terkecil.

METODOLOGI PENELITIAN

Data yang digunakan dalam penelitian ini adalah data Riset Kesehatan Dasar (RKD) Indonesia tahun 2007, khususnya data RKD Provinsi Papua tahun 2007 yang bersumber dari Badan Penelitian dan Pengembangan Kesehatan Departemen Kesehatan Republik Indonesia. Populasi dalam analisis ini adalah seluruh rumah tangga di Provinsi Papua. Sedangkan sampel (8) dengan memanfaatkan sampel RKD 2007. Variabel respon (Y) pada penelitian ini adalah jumlah penderita filariasis tiap kabupaten/kota di Provinsi Papua dengan jumlah pengamatan sebanyak 20. Sedangkan variabel prediktor (X) untuk penelitian ini adalah sebagai berikut :

1. Persentase rumah tangga yang tinggal di pedesaan untuk tiap kabupaten/kota di Provinsi Papua (X_1)
2. Persentase penduduk yang berjenis kelamin laki-laki untuk tiap kabupaten/ kota di Provinsi Papua (X_2)
3. Persentase penduduk yang berusia 20-39 tahun untuk tiap kabupaten/kota di Provinsi Papua (X_3)
4. Persentase penduduk yang tidur di dalam kelambu untuk tiap kabupaten/kota di Provinsi Papua (X_4)
5. Persentase penduduk yang tidur di dalam kelambu berinsektisida untuk tiap kabupaten/kota di Provinsi Papua (X_5)
6. Persentase rumah tangga yang menggunakan tempat penampungan air minum terbuka untuk air minum sebelum dimasak untuk tiap kabupaten/kota di Provinsi Papua (X_6)
7. Rata-rata jarak yang harus ditempuh ke sarana pelayanan kesehatan terdekat untuk tiap kabupaten/kota di Provinsi Papua (X_7)
8. Rata-rata waktu tempuh ke sarana pelayanan kesehatan terdekat untuk tiap kabupaten/kota di Provinsi Papua (X_8)
9. Persentase rumah tangga yang menggunakan racun serangga/pembasmi hama selama sebulan yang lalu untuk tiap kabupaten/kota di Provinsi Papua (X_9)
10. Presentase rumah tangga yang memelihara hewan peliharaan (anjing/kucing/kelinci) untuk tiap kabupaten/kota di Provinsi Papua (X_{10})

Berikut ini adalah langkah-langkah analisis data yang digunakan dalam penelitian ini.

1. Menentukan model regresi Poisson
2. Menaksir parameter model regresi Poisson
3. Menentukan devians (simpangan) model
4. Menentukan model terbaik regresi Poisson
   Model regresi Poisson yang layak digunakan dipilih berdasarkan nilai devians yang kecil.
5. Melakukan uji overdispersi
6. Menaksir paramater model regresi ZIP
7. Menguji kesesuaian model regresi ZIP
8. Menguji hipotesis model regresi ZIP
9. Menentukan model terbaik regresi ZIP

Tabel 1 Hipotesis dan Statistik Uji untuk Pengujian Parameter Regresi *Zero-Inflated* Poisson

No.	Pengujian	Hipotesis	Statistik Uji
1	Kesesuaian Model	$H_0: \beta_1 = \beta_2 = \ldots = \beta_r = \gamma_1 = \gamma_2 = \ldots = \gamma_r = 0$ H_1: paling sedikit ada satu $\beta_i \neq 0$ atau $\gamma_i \neq 0$	$G = \left(2\sum_{i=1}^{n} \left(z_i x_i^T \hat{\gamma} - \ln(1 + \exp(x_i^T \hat{\gamma}) \right) + 2\sum_{i=1}^{n} (1 - z_i) \left(y_i x_i^T \hat{\boldsymbol{\beta}} - \exp(x_i^T \hat{\boldsymbol{\beta}}) \right) \right)$ $- \left(2\sum_{i=1}^{n} z_i \hat{\gamma}_0 - \ln(1 + x_i^T \hat{\gamma}_0) + 2\sum_{i=1}^{n} (1 - z_i) \left(y_i \hat{\beta}_0 - \exp(\hat{\beta}_0) \right) \right)$
2	Individu (parameter model log)	$H_0: \beta_i = 0$ $H_1: \beta_i \neq 0$	$G = \left(2\sum_{i=1}^{n} \left(z_i x_i^T \hat{\gamma} - \ln(1 + \exp(x_i^T \hat{\gamma}) \right) + 2\sum_{i=1}^{n} (1 - z_i) \left(y_i x_i^T \hat{\boldsymbol{\beta}} - \exp(x_i^T \hat{\boldsymbol{\beta}}) \right) \right) +$ $- 2\sum_{i=1}^{n} (1 - z_i) \left(y_i x_i^T \hat{\beta}_i - \exp(x_i^T \hat{\beta}_i) \right)$
3	Individu (parameter model logit)	$H_0: \gamma_i = 0$ $H_1: \gamma_i \neq 0$	$G = \sum_{i=1}^{n} \left(z_i x_i^T \hat{\gamma} - \ln(1 + \exp(x_i^T \hat{\gamma}) \right) + 2\sum_{i=1}^{n} (1 - z_i) \left(y_i x_i^T \hat{\boldsymbol{\beta}} - \exp(x_i^T \hat{\boldsymbol{\beta}}\right.$ $- 2\sum_{i=1}^{n} (1 - z_i) \ln(y_i)! - 2\sum_{i=1}^{n} \left(z_i \hat{\gamma}_0 - \ln(1 + \exp(\hat{\gamma}_0) \right)$

ANALISIS DAN PEMBAHASAN

Karakteristik Penderita

Tabel 2 menunjukkan statistik deskriptif dari variabel respon (y) dan variabel prediktor (x) yang digunakan untuk pemodelan regresi Poisson. Rata-rata persentase rumah tangga yang tinggal di pedesaan adalah 82,72. Rentangan persentase penduduk yang berjenis kelamin laki-laki berada pada 44,23 sampai 53,42. Rata-rata persentase penduduk yang berusia 20-39 tahun adalah 31,35. Persentase penduduk yang tidur di dalam kelambu antara 2,56 sampai 89,64. Penduduk yang tidur di dalam kelambu belum tentu menggunakan kelambu yang berinsektisida. Hal ini ditunjukkan dengan nilai persentase penduduk yang tidur di dalam kelambu berinsektisida masih berada di bawah persentase penduduk yang tidur di dalam kelambu.

Tabel 2 Statistik Deskriptif

Variabel	Mean	Minimum	Maksimum
Y	0,95	0,00	11,00
X_1	82,72	9,17	100,00
X_2	49,81	44,23	53,42
X_3	31,35	22,45	40,11
X_4	39,93	2,56	89,64
X_5	21,50	0,11	80,74
X_6	32,24	6,79	74,02
X_7	3,80	0,87	9,49
X_8	51,46	12,10	156,07
X_9	24,78	0,00	63,30
X_{10}	29,24	1,71	62,79

Rata-rata persentase rumah tangga yang menggunakan tempat penampungan air minum terbuka untuk air minum sebelum dimasak adalah 32,24. Rentangan rata-rata jarak yang harus ditempuh ke sarana pelayanan kesehatan berada pada 870 m sampai 9,49 km dan rentangan rata-rata waktu tempuhnya berada pada 12 menit 6 detik sampai 2 jam 36 menit 4 detik. Rata-rata persentase rumah tangga yang menggunakan racun serangga/ pembasmi hama adalah 24,78 dan rata-rata persentase rumah tangga yang memelihara hewan peliharaan (anjing/kucing/ kelinci) adalah 29,24.

Model Regresi Poisson

Pengujian distribusi Poisson pada variabel respon membuktikan bahwa variabel respon berdistribusi Poisson. Nilai T untuk uji Kolmogorov-Smirnov adalah 0,263 lebih besar dari $w_{0,95} = 0,265$. Taraf signifikansi yang digunakan dalam pengujian distribusi dan pengujian selanjutnya adalah adalah $\alpha = 0,1$ karena penelitian ini merupakan penelitian sosial.

Setelah dilakukan pengujian distribusi Poisson, langkah selanjutnya adalah melakukan pembentukan model regresi Poisson. Nilai penduga parameter model regresi Poisson disajikan pada Tabel 3. Dari Tabel tersebut terlihat parameter model regresi Poisson yang signifikan pada $\alpha = 0,1$ hanya parameter β_{10} . Sehingga perlu dicari model regresi Poisson lain dengan lebih banyak variabel prediktor yang signifikan. Kombinasi yang bisa dibuat dengan menggunakan beberapa kelompok variabel prediktor tertera pada Tabel 4.

Tabel 3 Nilai Dugaan parameter Regresi Poisson

Parameter	Nilai Dugaan	SE	G_{hitung}	Nilai-p
β_0	-7,69	32,85	0,05	0,81
β_1	-0,15	0,14	1,17	0,28
β_2	0,04	0,64	0,00	0,96
β_3	0,45	0,32	2,06	0,15
β_4	0,12	0,11	1,19	0,28
β_5	-0,06	0,06	0,96	0,33
β_6	-0,08	0,09	0,89	0,35
β_7	-0,04	0,35	0,02	0,90
β_8	0,01	0,03	0,15	0,69
β_9	-0,14	0,16	0,78	0,38
β_{10}	0,15	0,07	4,56	**0,03**

Tabel 4 Kombinasi Model yang Bisa Dibuat

Jumlah Variabel	Jumlah Kombinasi
10 variabel	1023
9 variabel	511
8 variabel	255
7 variabel	127
6 variabel	63
5 variabel	31
4 variabel	15
3 variabel	7
2 variabel	3
1 variabel	1

Kombinasi yang masih bisa dibuat adalah kombinasi dengan lima variabel, sehingga dicari lima variabel yang parameternya signifikan secara individu dengan memodelkan variabel respon dengan variabel prediktor satu per satu. Nilai penduga parameter model regresi Poisson dengan satu variabel prediktor tertera pada Tabel 5.

Berdasarkan Tabel 5 hanya parameter β_r untuk model regresi Poisson dengan variabel X_8 dan X_{10} yang memiliki nilai-p kurang dari 0,1. Sehingga, dicari variabel lain yang memiliki nilai -p kecil untuk parameter β_r, yaitu variabel X_4, X_5, dan X_6. Dari 5 variabel prediktor yang digunakan untuk pembentukan model Regresi Poisson didapat 31 kemungkinan model seperti pada Tabel 4. Hasil pendugaan parameter untuk 31 kemungkinan model disajikan pada Tabel 6.

Berdasarkan model yang dibentuk, pada Tabel 6 diatas ada 14 model yang layak digunakan (tanda *) mengingat nilai devians untuk 14 model tersebut merupakan nilai devians yang kecil untuk tiap-tiap kelompok model. Terdapat lima kelompok model berdasarkan jumlah variabel prediktor yang dimasukkan ke dalam model, yaitu kelompok model dengan satu variabel prediktor, 2 variabel prediktor, 3 variabel prediktor, 4 variabel prediktor, dan 5 variabel prediktor. Model-model yang layak digunakan untuk pemodelan Regresi Poisson, nilai devians dan $\chi^2_{(n-k-1,\alpha)}$ selanjutnya ditunjukkan dalam Tabel 7.

Tabel 5 Nilai Dugaan Parameter Regresi Poisson untuk Satu Prediktor

No.	Var	Nilai Dugaan		Nilai-p	
		β_0	β_r	β_0	β_r
1	X_1	-1,54	0,02	0,62	0,62
2	X_2	-2,93	0,06	0,78	0,78
3	X_3	-3,22	0,10	0,29	0,28
4	X_4	0,67	-0,02	0,25	0,18*
5	X_5	0,47	-0,03	0,38	0,27*
6	X_6	1,06	-0,04	0,20	0,20*
7	X_7	0,14	-0,05	0,85	0,77
8	X_8	-1,00	0,01	0,21	0,09*
9	X_9	0,36	-0,02	0,58	0,45
10	X_{10}	-1,93	0,05	0,10	0,05*

Seluruh nilai devians untuk 14 model tersebut lebih besar dari nilai Chi-Square tabel, sehingga dapat disimpulkan bahwa 14 model tersebut layak digunakan. Model-model yang dipilih tersebut, kesemuanya menunjukkan adanya overdispersi karena nilai devians dibagi dengan derajat bebasnya lebih besar dari 1. Model regresi Poisson tidak memenuhi asumsi yaitu $E(y_i)$ sama dengan $Var(y_i)$ sehingga perlu digunakan model lain untuk mengetahui faktor-faktor yang mempengaruhi kejadian filariasis di Provinsi Papua. Model yang diusulkan adalah model regresi ZIP karena data yang digunakan memiliki banyak nilai nol.

Model Regresi *ZIP*

Hasil estimasi paramater model regresi *ZIP*, G_{hitung} , dan AIC adalah ditunjukkan pada Lampiran 1. Pengujian parameter secara serentak untuk semua alternatif model membuktikan bahwa semua model alternatif layak digunakan. Model dengan nilai AIC terkecil adalah model regresi ZIP dengan dua variabel prediktor, yaitu X_8 dan X_{10}. Namun, jika dilihat dari jumlah parameter yang signifikan, maka model ini tidak sesuai karena hanya dua parameter yang signifikan dari enam parameter yang dimiliki.

Tabel 6 Nilai dugaan parameter dan devians pada setiap kemungkinan model regresi Poisson

No.	Model	Nilai Dugaan Paramater						Devians
		β_0	β_4	β_5	β_6	β_8	β_{10}	
1.	$\exp(\beta_0 + \beta_4 x_{41i})$	0,637	-0,022					53,97
2.	$\exp(\beta_0 + \beta_5 x_{5i})$	0,457		-0,034				54,47
3.	$\exp(\beta_0 + \beta_6 x_{6i})$	1,031			-0,040			53,72
4.	$\exp(\beta_0 + \beta_8 x_{8i})$	-0,915				0,013		52,61
5.	$\exp(\beta_0 + \beta_{10} x_{10i})$	-1,761					0,048	48,51*
6.	$\exp(\beta_0 + \beta_4 x_{41i} + \beta_5 x_{5i})$	0,613	-0,014	-0,016				53,60
7.	$\exp(\beta_0 + \beta_4 x_{41i} + \beta_6 x_{6i})$	1,334	-0,018		-0,030			50,13
8.	$\exp(\beta_0 + \beta_4 x_{41i} + \beta_8 x_{8i})$	-0,336	-0,010			0,010		51,95
9.	$\exp(\beta_0 + \beta_4 x_{4i} + \beta_{10} x_{10i})$	-1,381	-0,039				0,072	30,32*
10.	$\exp(\beta_0 + \beta_5 x_{5i} + \beta_6 x_{6i})$	1,140		-0,026	-0,029			50,67
11.	$\exp(\beta_0 + \beta_5 x_{5i} + \beta_8 x_{8i})$	-0,443		-0,016		0,010		51,73
12	$\exp(\beta_0 + \beta_5 x_{5i} + \beta_{10} x_{10i})$	-1,687		-0,058			0,068	35,86*
13.	$\exp(\beta_0 + \beta_6 x_{6i} + \beta_8 x_{8i})$	0,131			-0,041	0,015		45,34
14.	$\exp(\beta_0 + \beta_6 x_{6i} + \beta_{10} x_{10i})$	-0,714			-0,041		0,049	42,51
15.	$\exp(\beta_0 + \beta_8 x_{8i} + \beta_{10} x_{10i})$	-3,973				0,022	0,068	33,42*
16.	$\exp(\beta_0 + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i})$	1,291	-0,013	-0,010	-0,029			49,97
17.	$\exp(\beta_0 + \beta_4 x_{41i} + \beta_5 x_{5i} + \beta_8 x_{8i})$	-0,336	-0,004	-0,013		0,009		51,69
18.	$\exp(\beta_0 + \beta_4 x_{41i} + \beta_5 x_{5i} + \beta_{10} x_{10i})$	-1,476	-0,031	-0,017			0,074	29,81*
19.	$\exp(\beta_0 + \beta_4 x_{41i} + \beta_6 x_{6i} + \beta_8 x_{8i})$	0,090	0,001		-0,042	0,015		45,33
20.	$\exp(\beta_0 + \beta_4 x_{41i} + \beta_6 x_{6i} + \beta_{10} x_{10i})$	-1,242	-0,038		-0,005		0,071	30,26*
21.	$\exp(\beta_0 + \beta_4 x_{4i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	-2,028	-0,030			0,007	0,070	29,87*
22.	$\exp(\beta_0 + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_8 x_{8i})$	0,164		-0,002	-0,041	0,014		45,32
23.	$\exp(\beta_0 + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_{10} x_{10i})$	-1,213		-0,048	-0,020		0,066	34,40*
24.	$\exp(\beta_0 + \beta_5 x_{5i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	-3,138		-0,029		0,015	0,070	31,50*
25.	$\exp(\beta_0 + \beta_6 x_{6i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	2,971			0,006	-0,009	0,017	62,04
26.	$\exp(\beta_0 + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_8 x_{8i})$	0,068	0,004	-0,005	-0,041	0,015		45,28
27.	$\exp(\beta_0 + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_{10} x_{10i})$	-1,368	-0,030	-0,017	-0,004		0,073	29,77*
28.	$\exp(\beta_0 + \beta_4 x_{41i} + \beta_5 x_{5i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	-1,982	-0,025	-0,014		0,006	0,072	29,53*
29.	$\exp(\beta_0 + \beta_4 x_{41i} + \beta_6 x_{6i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	-1,871	-0,024		-0,011	0,009	0,066	29,61*
20.	$\exp(\beta_0 + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	-2,441		-0,021	-0,021	0,015	0,064	30,24*
31.	$\exp(\beta_0 + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	-1,872	-0,012	-0,013	-0,010	0,007	0,068	29,34*

Tabel 7 Analisis Kesesuaian Model Regresi Poisson

No.	Model	Devians	*db*	$\chi^2_{(db,\alpha)}$	Devians/db
1.	$\mu_i = \exp(\beta_0 + \beta_{10} x_{10i})$	48,5098	18	25,989	2,6950
2.	$\mu_i = \exp(\beta_0 + \beta_4 x_{4i} + \beta_{10} x_{10i})$	30,3207	17	24,769	1,7836
3.	$\mu_i = \exp(\beta_0 + \beta_5 x_{5i} + \beta_{10} x_{10i})$	35,8680	17	24,769	2.1099
4.	$\mu_i = \exp(\beta_0 + \beta_8 x_{8i} + \beta_{10} x_{10i})$	33.4210	17	24,769	1.9659
5.	$\mu_i = \exp(\beta_0 + \beta_4 x_{41i} + \beta_5 x_{5i} + \beta_{10} x_{10i})$	29,8139	16	23,542	1,8634
6.	$\mu_i = \exp(\beta_0 + \beta_4 x_{41i} + \beta_6 x_{6i} + \beta_{10} x_{10i})$	30,2698	16	23,542	1,8919
7.	$\mu_i = \exp(\beta_0 + \beta_4 x_{4i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	29,8755	16	23,542	1,8672
8.	$\mu_i = \exp(\beta_0 + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_{10} x_{10i})$	34.4008	16	23,542	2.1500
9.	$\mu_i = \exp(\beta_0 + \beta_5 x_{5i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	31,4964	16	23,542	1,9685
10.	$\mu_i = \exp(\beta_0 + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_{10} x_{10i})$	29,7797	15	22,307	1,9853
11.	$\mu_i = \exp(\beta_0 + \beta_4 x_{41i} + \beta_5 x_{5i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	29,5356	15	22,307	1,9690
12	$\mu_i = \exp(\beta_0 + \beta_4 x_{41i} + \beta_6 x_{6i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	29,6177	15	22,307	1,9745
13.	$\mu_i = \exp(\beta_0 + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	30,2480	15	22,307	2,0165
14.	$\mu_i = \exp(\beta_0 + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_8 x_{8i} + \beta_{10} x_{10i})$	29,3395	14	21,064	2,0957

Model selanjutnya yang memiliki nilai AIC terkecil adalah model dengan tiga variabel prediktor, yaitu X_4, X_5, dan X_{10}. Parameter yang signifikan adalah γ_0, γ_4, γ_5, γ_{10}, dan β_{10}.

Tabel 8 merupakan nilai dugaan parameter jika parameter yang tidak signifikan dihilangkan. Jika beberapa parameter dikeluarkan maka banyak parameter yang tidak signifikan, sehingga model yang dipakai adalah model awal yaitu :

$\log(\mu_i) = 0{,}257 - 0{,}004X_{4i} - 0{,}068X_{5i} + 0{,}04X_{10i}$ dan

$\text{logit}(\omega_i) = 184{,}1 + 14{,}655X_{4i} - 44{,}066X_{5i} - 4{,}816X_{10i}$

dengan X_4 menyatakan persentase penduduk yang tidur di dalam kelambu, X_5 menyatakan persentase penduduk yang tidur di dalam kelambu berinsektisida, dan X_{10} menyatakan persentase rumah tangga yang memelihara hewan peliharaan (anjing/kucing/ kelinci). Model logit regresi ZIP menjelaskan bahwa peluang respon (y_i) bernilai nol dipengaruhi oleh persentase anggota rumah tangga yang tidur di dalam kelambu, persentase anggota rumah tangga yang tidur di dalam kelambu berinsektisida, dan persentase rumah tangga yang memeli-hara hewan peliharaan (anjing/kucing/kelinci). Sedangkan model log menjelaskan bahwa semakin besar persentase rumah tangga yang memelihara hewan peliharaan (anjing/kucing/kelinci) akan menaikkan jumlah penderita filariasis di Provinsi Papua. Kenaikan persentase rumah tang-ga yang memelihara hewan peliharaan (anjing/kucing/ kelinci) sebesar satu satuan akan mening-katkan jumlah penderita filariasis di Provinsi Papua sebanyak dua orang.

Setelah pemodelan dengan regresi ZIP, dihitung nilai μ_i, ω_i dan nilai probabilitas banyaknya penderita penyakit kaki gajah (filariasis) untuk tiap kabupaten/kota di Provinsi Papua ($P(Y_i = y_i)$). Probabilitas banyaknya penderita filariasis di kabupaten Asmat, kabupaten Tolikara, kabupaten Supiori, kabupaten Yapen Waropen, dan Kota Jayapura cukup tinggi sehingga kabupaten/kota tersebut perlu mendapatkan perhatian untuk menyukseskan program eliminasi kaki gajah.

Tabel 8 Nilai dugaan parameter dan uji kesesuaian model regresi *ZIP* untuk prediktor X_4, X_5, dan X_{10}

Par	Estimasi	DF	t	*Nilai-p*	G_{hitung}
γ_0	0,8738	20	0,70	0,49	
γ_4	0,0030	20	0,10	0,92	
γ_5	-0,0010	20	-0,03	0,98	55,1
γ_{10}	-0,0186	20	-0,53	0,60	
β_{10}	0,0278	20	4,99	<0,0001	

Tabel 9 Probabilitas Banyaknya Penderita Penyakit Kaki Gajah (Filariasis) untuk tiap Kabupaten/ Kota di Provinsi Papua

Nama Kabupaten/Kota	μ_i	ω_i	$P(Y_i=0)$	$P(Y_i=1)$	$P(Y_i=2)$	$P(Y_i=3)$	$P(Y_i=4)$	$P(Y_i=5)$
1 .Merauke	3,693	0	0,025	0,092	0,170	0,209	0,193	0,143
2 .Jayawijaya	7,125	2,931E-09	0,001	0,006	0,020	0,049	0,086	0,123
3 .Jayapura	3,577	0	0,028	0,100	0,179	0,213	0,191	0,136
4 .Nabire	3,476	5,506E-71	0,031	0,108	0,187	0,217	0,188	0,131
5 .Puncak Jaya	5,773	1	1	0	0	0	0	0
6. Mimika	4,971	0	0,007	0,035	0,086	0,142	0,177	0,176
7. Boven Digoel	1,316	1	1	0	0	0	0	0
8. Mappi	2,247	1	1	0	0	0	0	0
9. Asmat	1,855	1,609E-104	0,157	0,290	0,269	0,166	0,077	0,029
10. Yahukimo	4,801	7,072E-66	0,008	0,040	0,095	0,152	0,182	0,175
11. Peg. Bintang	9,439	3,501E-33	0,000	0,001	0,004	0,011	0,026	0,050
12. Tolikara	2,297	3,169E-77	0,101	0,231	0,265	0,203	0,117	0,054
13. Sarmi	1,071	1	1	0	0	0	0	0
14. Keerom	2,031	1	1	0	0	0	0	0
15 Waropen	2,290	1	1	0	0	0	0	0
16 .Supiori	2,239	5,775E-07	0,107	0,239	0,267	0,199	0,112	0,050
17. Yapen Waropen	1,419	0	0,242	0,343	0,244	0,115	0,041	0,012
18. Biak Numfor	6,378	1	1	0	0	0	0	0
19. Paniai	12,325	0	0	0,000	0,000	0,001	0,004	0,011
20. Kota Jayapura	2,162	1,053E-19	0,115	0,249	0,269	0,194	0,105	0,045

KESIMPULAN DAN SARAN

Kesimpulan

1. Rata-rata persentase rumah tangga yang tinggal di pedesaan adalah 82,72. Rentangan persentase penduduk laki-laki berada pada 44,23- 53,42. Rata-rata persentase penduduk yang berusia 20-39 tahun adalah 31,35. Persentase penduduk yang tidur di dalam kelambu antara 2,56 - 89,64. Rata-rata persentase rumah tangga yang menggunakan tempat penampungan air minum terbuka untuk air minum sebelum dimasak adalah 32,24. Rentangan rata-rata jarak yang harus ditempuh ke sarana pelayanan kesehatan berada pada 870 m sampai 9,49 km dan rentangan rata-rata waktu tempuhnya berkisar antara 12 menit 6 detik sampai 2 jam 36 menit 4 detik. Rata-rata persentase rumah tangga yang menggunakan racun serangga/pembasmi hama adalah 24,78 dan rata-rata persentase rumah tangga yang memelihara hewan peliharaan (anjing/kucing/kelinci) adalah 29,24.
2. Model regresi Poisson tidak memenuhi asumsi rata-rata sama dengan varians atau terjadi overdispersi pada model regresi Poisson sehingga perlu digunakan model lain untuk mengetahui faktor-faktor yang mempengaruhi kejadian filariasis di Provinsi Papua. Model yang diusulkan adalah model regresi *zero-inflated* Poisson (*ZIP*) karena data yang digunakan memiliki banyak data yang bernilai nol.
3. Model regresi *ZIP* terbaik adalah sebagai berikut :

$$\log(\mu_i) = 0{,}257 - 0{,}004X_{4i} - 0{,}068X_{5i} + 0{,}04X_{10i}$$

dan

$$\operatorname{logit}(\omega_i) = 184{,}1 + 14{,}655X_{4i} - 44{,}066X_{5i} - 4{,}816X_{10i}$$

dimana X_4 X_5 dan X_{10} menyatakan % penduduk yang tidur di dalam kelambu, % penduduk yang tidur di dalam kelambu berinsektisida, % rumah tangga yang memelihara hewan peliharaan (anjing/kucing-/kelinci). Model logit regresi *ZIP* menjelaskan bahwa peluang jumlah penderita filariasis di kabupaten/kota yang bernilai nol dipengaruhi oleh persentase penduduk yang tidur di dalam kelambu, persentase anggota rumah tangga yang tidur di dalam kelambu berinsektisida, dan persentase rumah tangga yang memelihara hewan peliharaan (anjing/ kucing/kelinci). Sedangkan model log menjelaskan bahwa semakin besar persentase rumah tangga yang memelihara hewan peliharaan (anjing/kucing/kelinci) sebanyak satu satuan akan meningkatkan jumlah penderita filariasis di Provinsi Papua sebanyak dua orang.

4. Kabupaten/kota di Provinsi Papua yang perlu mendapatkan perhatian khusus dalam program eliminasi filariasis adalah Asmat, Tolikara, Supiori, Yapen Waropen, dan Kota Jayapura.

Saran

Penelitian ini menggunakan jumlah pengamatan yang kecil, sehingga kurang memungkinkan untuk menggunakan banyak variabel prediktor. Selain itu, untuk pengujian yang berhubungan dengan distribusi Chi-Square seperti yang digunakan dalam penelitian ini seharusnya jumlah pengamatan banyak. Untuk penelitian selanjutnya, jumlah pengamatan hendaknya menjadi suatu pertimbangan.

DAFTAR PUSTAKA

Ambarita LP, Sitorus, H. 2006. Studi Komunitas Nyamuk di Desa Sebubus (Daerah Endemis Filariasis), Sumatera Selatan Tahun 2004. *Jurnal Ekologi Kesehatan* 5(1):368– 375.

Baharuddin. 2005. Ukuran R^2 dalam Model Regresi Poisson. *Integral* 10(3):114-121.

Cameron AC, Trivedi PK. 1998. *Regression Analysis of Count Data*. Cambridge : Cambridge University Press.

Dalrymple ML, Hudson IL, Ford RPK. 2002. Finite Mixture, Zero-inflated Poisson and Hurdle models with application to SIDS. *Computational Statistics & Data Analysis* 41:491-504.

Jansakul N, Hinde JP, 2001. Score Tests for Zero-Inflated Poisson Models. *Computational Statistics & Data Analysis* 40:75-96.

Khoshgoftaar TM, Gao K, Szabo RM. 2004. Comparing software fault predictions of pure and zero-inflated Poisson regression models. *International Journal of System Science* 36(11): 705-715.

Kleinbaum DG, Kupper LL, Muller KE. 1988. *Applied Regression Analysis and Other Multivariable Methods***.** 2nd edition. Boston : PWS-KENT Publishing Company.

Lestari A, 2008. Pemodelan Regresi Zero Inflated Poisson (Aplikasi Pada Data Pekerja Seks Komersial di Klinik Reproduksi Putat Jaya Surabaya), [Tesis] Surabaya : Program Studi Magister, Jurusan Statistika, Fakultas MIPA, Institut Teknologi Surabaya.

Myers RH. 1990. *Classical and Modern Regression with* Applications . 2nd edition Boston : PWS-KENT Publishing Company.

Lampiran 1 Nilai Estimasi Parameter , G_{hitung}, dan AIC Model regresi *ZIP*

No.	Variabel	Nilai Estimasi Paramater												G_{hitung}	*AIC*
		γ_0	γ_4	γ_5	γ_6	γ_8	γ_{10}	β_0	β_4	β_5	β_6	β_8	β_{10}		
1.	X_{10}	-1,3683					0,03101	0,06554					-1,8049	51,8	59,8
2.	X_4 dan X_{10}	-0,04285	-0,0142				0,01106	-0,8712	-0,02767				0,06497	42,1	54,1
3.	X_5 dan X_{10}	2,56380		-0,15580			-0,03274	0,34970		-0,07113			0,03769	41,1	53,1
4.	X_8 dan X_{10}	11,8682				3,8903	-13,4691	-2,3320				0,0371	0,0111	34,8	46,8*
5.	X_4, X_5, dan X_{10}	184,100	14,655	-44,066			-4,816	0,257	-0,004	-0,068			0,040	42,1	47,5*
6.	X_4, X_6, dan X_{10}	-13,0692	1,6209		-9,4491		2,7985	-1,2954	-0,0083	-0,0563			0,0905	35,4	51,4
7.	X_4, X_8, dan X_{10}	47,2951	-4,7695			5,9443	-14,3295	-1,4127	-0,0083			0,0316	0,0054	33,2	49,2
8.	X_5, X_6, dan X_{10}	2,39720		-0,17640	0,00739	-0,03333		0,80690		-0,05574	-0,0277		0,03777	37,8	55,9
9.	X_5, X_8, dan X_{10}	45,1767		-5,0411		-8,2973	-1.5457			-0,0111		0,0336	0,0029	33,1	49,1
10.	X_4, X_5, X_6, dan X_{10}	11,8321	4,8093	-17,1301	2,1098		-1,5171	-0,1148	0,0060	-0,0690	-0,0189		0,0526	34,0	54
11.	X_4, X_5, X_8, dan X_{10}	42,5137	3,8899	-14,2658		5,3390	-13,7988	-1,3591	-0,0042	-0,0087		0,0315	0,0042	33,1	53,1
12	X_4, X_6, X_8, dan X_{10}	0,86140	0,71520		-1,71410	1,45810	-2,95000	-1,03860	0,00019		-0,01766	0,03551	-0,00563	32,5	52,5
13.	X_4, X_6, X_8, dan X_{10}	2,26780		2,54490	-2,56750	2,77440	-5,77420	-0,84200		0,00139	-0,01844	0,03580	-0,01030	31,9	51,9
14.	X_4, X_5, X_6, X_8, dan X_{10}	0,3066	0,0000	-6,8651	-1,9509	6,7845	-14,6029	0,0125	-0,0006	-0,0104	-0,0244	0,0331	-0,0197	33,7	57,7